

Garantindo acervos para o futuro: Plano de preservação digital para o Repositório Institucional Arca

Andréa Gonçalves do Nascimento

Doutoranda em Ciência da Informação pela Universidade Federal do Rio de Janeiro (UFRJ) - RJ - Brasil.
Mestrado profissional em Biblioteconomia pela Universidade Federal do Estado do Rio de Janeiro (Unirio) – RJ - Brasil. Bibliotecária da Fundação Oswaldo Cruz, (Fiocruz) - Brasil.
<http://lattes.cnpq.br/2052432238666012>
E-mail: andrea.goncalves@icict.fiocruz.br

Claudete Fernandes de Queiroz

Mestrado profissional em História, Política e Bens Culturais pela Fundação Getúlio Vargas (FGV) - Brasil. Tecnologista em Saúde Pública – Bibliotecária da Fundação Oswaldo Cruz (Fiocruz) - Rio de Janeiro, RJ - Brasil.
<http://lattes.cnpq.br/5902232749593657>
E-mail: claudete.queiroz@icict.fiocruz.br

Luciana Danielli de Araújo

Mestrado em Ciência da Informação pelo Programa de Pós-graduação em Ciência da Informação (PPGCI) - convênio Universidade Federal do Rio de Janeiro e Instituto Brasileiro de Informação em Ciência e Tecnologia (UFRJ/Ibict) - Rio de Janeiro, RJ - Brasil. Tecnologista da Fundação Oswaldo Cruz - Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (Fiocruz/ICICT) - Rio de Janeiro, RJ - Brasil.
<http://lattes.cnpq.br/0698873941750070>
E-mail: luciana.danielli@icict.fiocruz.br

Data de submissão: 13/09/2019. Data de aprovação no ConfOA: 12/06/2019. Data de publicação:

RESUMO

Este trabalho descreve o processo de construção do Plano de Ação de Preservação Digital para o Arca – Repositório Institucional da Fiocruz. A implementação de ações para a preservação digital ainda encontra desafios consideráveis em todos os seus estágios, sendo o primeiro a formulação de políticas institucionais e de outros instrumentos normativos e processuais necessários à sua execução. Um plano de preservação digital tem como objetivo sistematizar as etapas e procedimentos necessários para atender aos requisitos mínimos para a preservação dos objetos digitais depositados no repositório institucional. A construção do Plano de Ação foi composta por quatro etapas: análise dos requisitos, diretrizes e normas nacionais e internacionais que orientam as ações de preservação digital; definição dos elementos e da estrutura do documento; diagnóstico do estado atual do repositório Arca em relação aos elementos mínimos desejados; e definição das estratégias a serem adotadas para o repositório Arca. Como resultado, o Plano de Ação de Preservação Digital do repositório Arca estabelece os padrões que visam garantir que a produção científica da Fiocruz ali depositada seja preservada de modo permanente, em um ambiente confiável e seguro, buscando ainda contribuir para o debate teórico e prático sobre a preservação digital em repositórios.

Palavras-chave: Preservação digital. Repositório institucional. Plano de ação.

Securing collections for the future: a digital preservation plan for Arca Institutional Repository

ABSTRACT

This paper presents the process of building the Digital Preservation Action Plan for Fiocruz Institutional Repository – Arca. Digital preservation actions still faces considerable challenges in all its stages, the first being the formulation of institutional policies and other normative and procedural instruments needed for its implementation. A digital preservation plan aims to systematize the steps and procedures necessary to meet the minimum requirements for the preservation of digital objects in institutional repository. The Action Plan was developed following four steps: analysis of national and international guidelines and standards that guide digital preservation actions; definition of elements and structure of the document; diagnosis of the current state of the Arca repository in relation to the desired minimum elements; and definition of strategies to be adopted for the Arca repository. As a result, the Action Plan sets the standards that aim to ensure that scientific outputs by Fiocruz are permanently preserved, in a reliable and safe environment, and seeks to contribute to the theoretical and practical debate about digital preservation in repositories.

Keywords: Digital preservation. Institutional repository. Action Plan.

Asegurando colecciones para el futuro: plan de preservación digital para el Repositorio Institucional Arca

RESUMEN

Describe el proceso de construcción del Plan de Acción de Preservación Digital para el Depósito Institucional de Fiocruz – Arca. La implementación de acciones para la preservación digital aún enfrenta desafíos considerables en todas sus etapas, siendo la primera la formulación de políticas institucionales y otros instrumentos normativos y procesales. Un plan de preservación digital tiene como objetivo sistematizar los pasos y procedimientos necesarios para cumplir con los requisitos mínimos para la preservación de objetos digitales depositados en el repositorio institucional. La construcción del Plan de Acción se compuso de cuatro pasos: análisis de los requisitos, directrices y estándares nacionales e internacionales que guían las acciones de preservación digital; definición de elementos y estructura del documento; diagnóstico del estado actual del repositorio Arca en relación con los elementos mínimos deseados; y definición de estrategias a adoptar para el repositorio de Arca. Como resultado, el Plan de Acción establece los estándares que tienen como objetivo garantizar que la producción científica de Fiocruz se conserve permanentemente en un entorno confiable y seguro, y busca contribuir al debate teórico y práctico sobre preservación digital en repositorios.

Palabras clave: Preservación digital. Repositorio institucional. Plan de acción.

INTRODUÇÃO

Os repositórios institucionais já se consolidaram nas instituições de pesquisa brasileiras como importante ferramenta para a produção e circulação do conhecimento, tendo o papel de reunir, organizar, disseminar e preservar a produção científica dessas instituições. Entre suas funções, a de preservação ainda é a que requer mais atenção e ações concretas por parte dos gestores de repositórios. Apesar de a preservação digital ser um tema bastante discutido internacionalmente, sua efetiva implementação ainda encontra desafios consideráveis em todas as suas etapas, a começar pela formulação de políticas institucionais e outros instrumentos normativos e processuais necessários à sua execução.

A preservação digital constitui-se de um conjunto de ações para assegurar a longevidade dos documentos digitais (BULLOCK, 1999). Na Carta para Preservação do Patrimônio Arquivístico Digital (CONARQ, 2005), o Conselho Nacional de Arquivos afirma que essas ações devem ser incorporadas em todo o ciclo de vida dos documentos, desde a criação até o armazenamento e acesso, a fim de que não haja perda ou adulteração dos registros, e assim garantir que permaneçam disponíveis, recuperáveis e compreensíveis pelo tempo que se fizer necessário.

As ações de preservação digital se aplicam tanto a documentos criados originalmente no formato digital, chamados nato-digitais, como a produtos da conversão do formato impresso ou analógico para digital, conhecidos como representantes digitais (BULLOCK, 1999) e devem abranger todas as características essenciais do objeto digital, incluindo suas dimensões física, lógica e intelectual. A dimensão física diz respeito aos suportes de armazenamento digital, como dispositivos e mídias ópticas e magnéticas, para os quais devem ser previstas soluções relativas à compatibilidade, manutenção das mídias e transferência ou rejuvenescimento das mesmas, caso necessário. A dimensão lógica trata dos formatos de arquivo, sua interpretação por hardwares e softwares e a necessidade de atualização dos formatos para

garantir a continuidade da leitura e interpretação da cadeia de bits. Já a dimensão intelectual abarca todo o conteúdo informacional do objeto digital e os mecanismos que asseguram sua integridade e autenticidade, bem como a preservação de aspectos da forma de apresentação original do documento (ARELLANO, 2004; SANTOS, FLORES, 2017).

A preservação digital nos repositórios institucionais é uma questão de elevada importância e a instituição que os abriga tem um papel crucial, pois deve ser capaz de gerir toda sua informação digital a fim de mantê-la estável, utilizável e confiável, garantindo assim os seus propósitos operacionais atuais e futuros.

Santos e Flores (2017) reconhecem que a preservação de documentos digitais envolve tanto os procedimentos técnicos destinados a manter a autenticidade, a confiabilidade e a acessibilidade dos acervos digitais ao longo do tempo, como também os aspectos políticos e administrativos que visam garantir a longevidade dos documentos a partir da implementação de instruções normativas e processuais.

Do ponto de vista técnico-conceitual, o modelo de referência adotado internacionalmente é o Open Archival Information System (OAIS), disposto na norma ISO 14721:2003 e adaptado no Brasil pela norma NBR 15472:2007 (Modelo de referência para um sistema aberto de arquivamento de informação – SAAI). O modelo de referência OAIS estabelece uma estrutura terminológica e conceitual que visa facilitar o entendimento, a descrição e a comparação de estratégias e técnicas para preservação e acesso de longo prazo a arquivos digitais, e é aplicável a objetos digitais depositados em repositórios institucionais. O OAIS define um modelo de configuração e operação de um repositório digital confiável e descreve a estrutura e o fluxo das informações nesse sistema, desde o momento da inserção dos objetos digitais e metadados no sistema, até o seu armazenamento e acesso pelo consumidor final. No entanto, o modelo OAIS não determina de que forma isso deve ser implementado (em um software de sistema de preservação, por exemplo),

deixando a cargo de fornecedores e desenvolvedores de sistemas a concretização dos aspectos operacionais. Porém, entre as responsabilidades obrigatórias para a operação de um sistema de arquivamento de informações sob o modelo OAIS está a documentação das políticas e procedimentos adotados para garantir a preservação dos objetos digitais em longo prazo, incluindo um plano de sucessão, caso o arquivo seja desativado.

Neste trabalho, apresentamos o processo de desenvolvimento do Plano de Ação de Preservação Digital para o Arca – Repositório Institucional da Fiocruz, descrevendo as etapas que envolveram sua preparação e racionalizando a adoção de padrões, procedimentos e tecnologias para garantir a preservação, o acesso e a utilização dos seus acervos digitais no futuro.

O REPOSITÓRIO INSTITUCIONAL ARCA E A POLÍTICA DE PRESERVAÇÃO DA FIOCRUZ

O Arca é o repositório institucional da Fundação Oswaldo Cruz (Fiocruz)¹ e tem como função reunir, hospedar, disponibilizar e dar visibilidade à produção intelectual da instituição, estimulando a ampla circulação do conhecimento, fortalecendo o compromisso institucional com o livre acesso à informação, além de conferir transparência e incentivar a comunicação científica entre os pesquisadores, educadores, acadêmicos e gestores das áreas da saúde e destes com a sociedade civil.

O Arca foi criado em 2007, utilizando o software livre Dspace, e lançado oficialmente como repositório institucional em 2011. Em 2014, a Fiocruz instituiu a Política de Acesso Aberto ao Conhecimento, que “visa garantir à sociedade o acesso gratuito, público e aberto ao conteúdo integral de toda obra intelectual produzida pela Fiocruz” (FUNDAÇÃO OSWALDO CRUZ, 2014). Entre outros aspectos, a política determina o caráter mandatório do depósito no repositório Arca das dissertações e teses defendidas nos programas de pós-graduação da Fiocruz e dos artigos científicos publicados em periódicos de autoria de seus colaboradores.

O repositório Arca é mantido pelo Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT/Fiocruz) e seu acervo digital está organizado em comunidades, que correspondem às unidades da Fiocruz, que por sua vez se organizam em coleções que abrigam diferentes tipologias de documentos. Em respeito à diversidade de tipologias e suportes informacionais encontrados na produção científica da instituição, optou-se por estimular o autoarquivamento por parte da comunidade acadêmica, disponível através de login e senha no site do repositório. Assim, cada pesquisador é habilitado a realizar o depósito de sua produção acadêmica em sua respectiva Comunidade e Coleção, autorizando o acesso e uso da obra, de acordo com os termos de cessão não exclusiva de direitos autorais, sendo responsável pelo preenchimento dos metadados básicos de identificação do trabalho. A validação do autoarquivamento é realizada pelas bibliotecas físicas que compõem a Rede de Bibliotecas da Fiocruz, como instância responsável pelo processamento, disseminação e preservação das coleções.

Em consonância com as ações e práticas estabelecidas na política de acesso aberto, a Fiocruz lançou em 2018 sua Política de Preservação dos Acervos Científicos e Culturais, com a finalidade de preservar, valorizar e divulgar o patrimônio histórico, cultural e científico institucional e contribuir para a preservação da memória das áreas de Ciências da Saúde e Biomédicas (FUNDAÇÃO OSWALDO CRUZ, 2018). A Política de Preservação dos Acervos estabelece os princípios e diretrizes orientadores das ações de preservação a serem desenvolvidas para os diferentes acervos e coleções científicos e culturais da instituição, incluindo a dimensão da preservação e acesso digital. No entanto, esse documento não trata dos procedimentos e práticas para a operacionalização das ações de preservação. Sobre esse aspecto, ressaltamos que:

¹ Disponível em: <http://www.arca.fiocruz.br>. Acesso em: 19 ago. 2019

Uma importante distinção deve ser feita entre *planos* de preservação concretos e *políticas* de alto nível que geralmente são feitas em um nível institucional e regulam restrições e estratégias fundamentais (BECKER *et al.*, 2009, tradução nossa).

Segundo os autores, o plano de preservação digital possui uma abordagem mais concreta e específica, e configura um verdadeiro plano de ação para garantir as condições mínimas de preservação de determinado conjunto de objetos digitais sob a guarda de uma instituição específica. Na definição adotada pelo projeto PLANETS²:

O plano de preservação leva em consideração as políticas de preservação, obrigações legais, restrições organizacionais e técnicas, requisitos de usuários e metas de preservação e descreve o contexto de preservação, as estratégias de preservação avaliadas e a decisão resultante de uma estratégia, incluindo o raciocínio da decisão. Também especifica uma série de etapas ou ações (chamadas de *plano de ação de preservação*), juntamente com responsabilidades, regras e condições para execução na coleção. Desde que as ações e sua implantação, bem como o ambiente técnico, permitam isso, esse plano de ação é uma definição de fluxo de trabalho executável (PLANETS, 2006, tradução nossa).

O principal objetivo do Plano de Ação de Preservação Digital para o repositório institucional Arca é sistematizar e documentar as orientações e procedimentos técnicos necessários para garantir os requisitos mínimos para preservação das coleções de documentos técnico-científicos depositados. O Plano de Ação é orientado tanto para a equipe responsável pela operação do repositório, como para as equipes dos núcleos das unidades institucionais³ responsáveis pela coleta, encaminhamento e inclusão de documentos digitais.

A elaboração de um plano de ação concreto para a preservação digital das coleções de documentos digitais do repositório Arca exigiu extenso esforço

de investigação, dado que as evidências encontradas inicialmente demonstraram a inexistência desse tipo de documentação em outras instituições nacionais que possuem repositórios.

O resultado de um diagnóstico realizado em 2018 pela Rede Sudeste de Repositórios Institucionais (RIAA/Sudeste)⁴ apresentou um dado relevante: nenhuma das 55 instituições participantes da Rede possui um plano de preservação digital para o repositório institucional. Em levantamento anterior realizado por Weitzel e Mesquita (2015), oito entre nove repositórios institucionais pesquisados na Região Sudeste não possuíam sequer uma política de preservação digital formalizada. Outra pesquisa feita em 38 universidades federais brasileiras (SILVA JUNIOR, 2017) apontou que nenhuma delas possui uma política de preservação digital para seus repositórios. Esses dados corroboraram a relevância do assunto e a urgente necessidade do desenvolvimento do Plano de Ação de Preservação Digital para o repositório institucional Arca.

ELABORAÇÃO DO PLANO DE AÇÃO

A elaboração do Plano de Ação de Preservação Digital para o repositório Arca foi composta por quatro etapas. Na primeira etapa, analisamos diversas diretrizes e recomendações das normas nacionais e internacionais que orientam as ações de preservação digital, como a ISO 14721:2003 (OAIS - Open Archives Information System) e NBR 15472:2007 (SAAI - Sistema Aberto de Arquivamento de Informação), além da consulta à literatura nacional e internacional a respeito da elaboração de planos de ação para preservação digital (BECKER, 2009; CORRADO; SANDY, 2017; DIGITAL PRESERVATION COALITION, 2018; FIVE COLLEGE CONSORTIUM, 2017; ISO, 2010), e exemplos de políticas e planos

² PLANETS (Preservation and Long-term Access through Networked Services) foi um projeto cofinanciado pela União Europeia para enfrentar os principais desafios da preservação digital. Para mais informações, consultar: <https://www.planets-project.eu/>

³ Os Núcleos de Acesso Aberto ao Conhecimento (NAACs) são a instância responsável pela coordenação, gestão, operação, participação, promoção e acompanhamento da adesão ao Repositório Institucional Arca, no âmbito de cada unidade.

⁴ A Rede Sudeste de Repositórios Institucionais faz parte da Rede Nacional de Repositórios coordenada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), que tem como objetivo promover o compartilhamento de informações e experiências entre seus membros. Atualmente a Rede Sudeste é composta por 55 instituições públicas e privadas de ensino e pesquisa da região.

de preservação digital em âmbito internacional (BRITISH LIBRARY, 2017; COLOMBIA, 2019; NATIONAL RECORDS OF SCOTLAND, 2014; PURDUE UNIVERSITY LIBRARIES, 2012; THE ROYAL LIBRARY, 2014).

Na segunda etapa, identificamos um conjunto de elementos essenciais e itens desejáveis que devem ser contemplados no Plano de Ação de Preservação Digital (tabela 1).

Tabela 1 – Elementos essenciais e itens desejáveis do Plano de Ação de Preservação Digital

Elementos	Itens desejáveis
Descrição do cenário institucional	Políticas de preservação relacionadas Obrigações legais Contexto de preservação
Descrição da coleção	Identificação e escopo Restrições organizacionais Restrições técnicas Requisitos do usuário
Avaliação de riscos	Possíveis ameaças externas Probabilidade e extensão do risco Plano para mitigar os riscos
Planejamento de preservação	Metas de preservação Estratégias de preservação avaliadas Decisão resultante, incluindo o raciocínio da decisão
Ações de preservação	Seleção de componentes para preservação Integridade e autenticidade do conteúdo Controle de formatos de arquivo Requisitos mínimos de metadados Gestão de direitos Responsabilidades sobre as etapas ou ações Regras e condições para execução na coleção
Plano de sucessão	Recriação de metadados e arquivos Planejamento de migração

Fonte: Elaborado pelas autoras.

A terceira etapa foi orientada pelo resultado de dois diagnósticos do estado atual do repositório Arca, em relação a requisitos de confiabilidade e de adequação aos níveis de preservação digital.

No primeiro diagnóstico, Lanzelotte *et al.* (2019) realizaram uma avaliação dos requisitos de confiabilidade no repositório institucional Arca, utilizando como referência uma lista de requisitos comuns para a avaliação de confiabilidade de repositórios digitais com base no conjunto de critérios adotados por três documentos: o padrão de certificação *Trustworthy Repositories Audit and Certification: Criteria and Checklist* (mais conhecida como TRAC), a norma ISO 16363:2012 (*Audit and Certification of Trustworthy Digital Repositories*) e as diretrizes nacionais para a implementação de repositórios digitais confiáveis de documentos arquivísticos, do Conselho Nacional de Arquivos (Conarq).

Como resultado, verificou-se que o Arca atende em algum grau a 63% dos critérios de infraestrutura organizacional, 73% dos critérios de gestão de objetos digitais e 50% dos requisitos de infraestrutura tecnológica e de segurança. No entanto, um dos pontos destacados pelos autores foi a falta de documentação dos procedimentos em geral, e especificamente, de um plano de preservação digital, que seria um dos principais instrumentos para garantir a confiança e a segurança de um repositório digital. Um resumo das recomendações derivadas do diagnóstico de confiabilidade é apresentado na tabela 2.

Tabela 2 – Recomendações do diagnóstico de confiabilidade do Arca

Categoria	Recomendação
Infraestrutura Organizacional	Realizar ações de auditoria e manutenção das políticas em vigor.
	Tornar transparente o planejamento financeiro de curto e longo prazo, sobretudo quanto ao equilíbrio adequado de risco, benefício, investimento e despesa.
	Explicitar a forma de suporte técnico ao repositório.
	Especificar claramente as informações de preservação que devem estar associadas aos documentos no momento da submissão.
	Utilizar identificadores únicos e visíveis em todos os registros, de acordo com padrões reconhecidos.
	Tornar transparentes os procedimentos de adesão, os registros de processamento interno e de auditoria.
	Instituir estratégias sólidas, atuais e documentadas de preservação, implementada e continuada, contendo os mecanismos de atualização dessas políticas, com alertas sobre obsolescência iminente.
	Documentar eventos de migração, transformação, cópia, armazenamento distribuído e histórico de processamento que possam afetar a confiança na preservação.
	Adotar algum mecanismo automatizado para identificar ameaças de segurança ou falhas no sistema de gerenciamento de acesso.
Infraestrutura Tecnológica e de Segurança	Adotar e documentar procedimentos para garantir a confiabilidade, segurança e análise de riscos da infraestrutura do sistema.
	Identificar e documentar o número de cópias e sincronizações de todos os objetos digitais armazenados, bem como a sua localização, escrita sem ambiguidade.
	Estabelecer e documentar requisitos de segurança para recuperação de dados em caso de desastres e comprometimento do sistema.
	Incluir nos requisitos de segurança quem são as pessoas responsáveis pelas ações. Incluir uma cópia do backup fora do local onde a informação é preservada

Fonte: Adaptado pelas autoras a partir de LANZELOTTE *et al.* (2019).

O segundo diagnóstico foi realizado pela Asociación Iberoamericana de Preservación Digital (Apredig)⁵ com base nos requisitos de Níveis de Preservação Digital da National Digital Stewardship Alliance (NDSA)⁶, um conjunto de recomendações que permite que as instituições avaliem o nível alcançado na preservação digital de materiais específicos sob sua custódia ou em toda

a sua infraestrutura de preservação (tabela 3). A NDSA sugere uma abordagem gradual de esforços para a criação de melhores práticas de preservação digital, organizados em cinco áreas funcionais: armazenamento e localização geográfica; correção de arquivos e integridade de dados; segurança das informações; metadados; e formatos de arquivo. A realização desse diagnóstico é importante para identificar os pontos de atenção e os passos iniciais necessários para aprimorar a prática da preservação digital aplicada a uma coleção ou instituição.

⁵ Uma ferramenta de solicitação de diagnóstico baseado nos Níveis de Preservação Digital da NDSA está disponível no site da Asociación Iberoamericana de Preservación Digital (Apredig), em: <http://www.apredig.org/>

⁶ NDSA Digital Preservation Levels, disponível em: <https://ndsas.org/activities/levels-of-digital-preservation/>

Tabela 3 – Diagnóstico do repositório Arca com base nos Níveis de Preservação Digital da NDSA

Área funcional	Resultado
Armazenamento e localização geográfica	O número de backups deve ser aumentado. Distribuir cópias de segurança em diferentes localizações geográficas (2 a 3 cópias desejáveis). Estabelecer controle periódico automatizado de estado dos sistemas de armazenamento e suportes físicos para controle da obsolescência.
Correção de arquivos e integridade de dados	Criar informações de integridade do conteúdo. Bloquear processos de gravação em mídia externa para evitar alterações não intencionais do arquivo. Monitorar a integridade do conteúdo em intervalos regulares. Confirmar que ninguém tenha acesso de gravação às cópias do conteúdo.
Segurança das informações	O diagnóstico não é totalmente aplicável, porque níveis anteriores devem ser cobertos. Recomendável garantir que as permissões de acesso da equipe ao conteúdo sejam suficientemente restritas. Aumentar o registro documentado de ações de gerenciamento de arquivos para melhorar a capacidade de segurança e a possibilidade de auditar possíveis incidentes.
Metadados	O diagnóstico não é totalmente aplicável, porque níveis anteriores devem ser cobertos. É essencial ter um inventário do conteúdo e dos locais de armazenamento, além de registrar metadados administrativos e de transformação para documentar incidentes. É necessário manter uma cópia separada de metadados do inventário.
Formatos de arquivo	É necessário ter um inventário dos formatos utilizados para avaliar e controlar o risco de obsolescência no curto prazo. Avaliar periodicamente a necessidade de executar migrações de formato de arquivo.

Fonte: Adaptado pelas autoras a partir de APREDIG, 2019.

A última etapa conclui com a definição das estratégias a serem adotadas especificamente para o repositório institucional Arca, considerando os elementos essenciais da estrutura do Plano de Ação e o resultado dos diagnósticos de confiabilidade e níveis de preservação.

ANÁLISE E DISCUSSÃO DOS RESULTADOS

A operação do modelo OAIS se baseia no conceito de pacote de informação (Information Package) como unidade de informação que reúne os conteúdos e metadados para a preservação do arquivo digital. De acordo com o modelo, o produtor deve submeter Pacotes de Informação para Submissão (Submission Information Package - SIP) ao sistema de preservação onde os metadados e os arquivos serão armazenados na forma de Pacotes de Informação para Arquivamento (Archival Information Package - AIP) e o sistema deve prover aos usuários acesso à informação que for solicitada através de um ou mais Pacotes de Informação para Disseminação (Dissemination Information Package - DIP).

No entanto, devemos considerar que a Fiocruz atualmente não possui um sistema de preservação digital em operação apto a receber pacotes de informação sob as diretrizes do modelo OAIS. Assim, o foco do Plano de Ação de Preservação do Arca é assegurar a preparação adequada dos conteúdos digitais para futuro ingresso em um sistema compatível com o modelo OAIS, porém garantindo a manutenção atual da integridade, confiabilidade e segurança desses conteúdos.

Com base nos itens elencados na tabela 1, foi elaborada a estrutura básica do Plano de Ação de Preservação Digital, composta de duas partes. A primeira parte descreve os elementos essenciais que norteiam as ações de preservação, como o cenário institucional, a descrição da coleção, a avaliação de riscos e ameaças, e o planejamento das estratégias para mitigação da obsolescência. Em relação ao último elemento, o repositório Arca optou por uma combinação de estratégias de preservação a serem

aplicadas em diferentes estágios do ciclo de vida do conteúdo digital, a exemplo da estratégia adotada pelo Arquivo Nacional⁷, sejam elas: no ingresso dos documentos no repositório, proceder a normalização de formatos de arquivo sob risco para um padrão de formato adequado à preservação; periodicamente, realizar o monitoramento dos formatos em uso a fim de detectar formatos que possam estar sob risco iminente; e, quando necessário, efetuar a migração de formatos de arquivo e/ou versão de softwares, a fim de mitigar a obsolescência tecnológica.

A segunda parte do documento descreve as estratégias escolhidas pelo repositório Arca para atendimento às ações de preservação, considerando o estado atual do repositório, as restrições institucionais e técnicas (tabela 4). As estratégias definidas serão detalhadas no Plano de Ação e devem ser revistas periodicamente, a fim de refletir quaisquer mudanças no estado e no fluxo de operação do repositório, bem como nas condições técnicas e institucionais.

Tabela 4 – Estratégias para as ações de preservação digital do repositório Arca

Ação de preservação	Estado atual e restrições	Estratégia
Seleção de componentes para preservação	Não possui inventário dos objetos digitais Não há critérios de seleção para o que deve ser preservado	Identificar os objetos digitais armazenados e seus componentes que devem ser preservados
Integridade e autenticidade do conteúdo	Não possui sistema de preservação digital ou outro sistema de criação e verificação de integridade e autenticidade Requisitos de segurança para acesso aos arquivos não estão documentados	Criar checksum para os objetos digitais preserváveis no momento da inserção no repositório e armazenar o valor. Restringir o acesso de gravação às cópias de conteúdo e documentar os requisitos de segurança e acesso ao conteúdo.
Controle de formatos de arquivo	Possui inventário de formatos depositados no repositório Não realiza controle periódico de obsolescência Formatos de arquivo submetidos ao Arca devem seguir as diretrizes do Manual de Digitalização.	Identificar e documentar os formatos atualmente em uso Definir os formatos de arquivo aceitos e recomendados como padrão para preservação Normalizar os formatos de arquivos no momento da inserção dos documentos no repositório Avaliar periodicamente a necessidade de executar migrações de formato de arquivo, a partir de alertas de obsolescência tecnológica
(Continua)		

⁷ Cf. ARQUIVO NACIONAL, 2016.

Tabela 4 – Estratégias para as ações de preservação digital do repositório Arca		(Conclusão)
Ação de preservação	Estado atual e restrições	Estratégia
Requisitos mínimos de metadados	Utiliza somente o padrão Dublin Core para metadados descritivos Metadados mínimos exigidos: Título, Ano de Publicação e Autor Não possui conteúdo de metadados de preservação	Inventariar atuais padrões e conteúdo de metadados Determinar estrutura mínima de metadados descritivos e de preservação, de acordo com padrões internacionais Especificar as informações de preservação que devem estar associadas ao objeto digital no momento da submissão Avaliar a capacidade dos metadados mínimos serem coletados via autoarquivamento ou processamento pelos Núcleos e bibliotecas
Gestão de direitos	Possui termo de cessão de direitos, assinada por todos os autores	Termo de cessão de direitos que permita a realização das atividades de preservação digital pelo repositório

Fonte: Elaborado pelas autoras.

A primeira ação de preservação corresponde à definição dos componentes do acervo que serão destinados à preservação, e inclui a definição dos critérios para a seleção e priorização dos objetos digitais e metadados que serão preservados (por tipo, prioridade e grau de risco de perda) e a definição do conjunto de propriedades significativas dos objetos digitais que devem ser preservados. As ações descritas nessa etapa preveem a preparação dos arquivos digitais e metadados para futuro ingresso no sistema de preservação.

A integridade e autenticidade do conteúdo não podem ser adequadamente controladas sem o uso de ferramentas específicas, que atualmente não são utilizadas pelo repositório. Inicialmente, será adotada a criação e armazenamento de valores de *checksum* para os objetos digitais, o que possibilitará identificar no futuro se foi produzida qualquer modificação na cadeia de bits do arquivo após ter sido inserido no repositório. Também serão reforçadas as restrições de acesso às cópias de conteúdo, a fim de minimizar a possibilidade de alteração de qualquer informação no conteúdo ou descrição do objeto digital.

O repositório Arca segue as especificações de formatos e metadados definidos em seu Plano Operativo e no Manual de Digitalização da Fundação Oswaldo Cruz. Como parte do Plano de Ação, os formatos de arquivo e metadados serão inventariados e documentados para permitir a posterior monitoração quanto à obsolescência, que indique a necessidade de migrar os arquivos para novos formatos ou versões. Outro ponto importante é a definição de um conjunto mínimo de metadados de preservação, que deverão ser especificados e submetidos juntamente com o objeto digital. Essa mudança no fluxo de operação demanda o envolvimento e treinamento das equipes das bibliotecas físicas e dos núcleos que contribuem para a alimentação de dados no repositório.

Quanto à gestão de direitos, o repositório Arca já conta com um termo de cessão que deve ser aceito pelos detentores dos direitos autorais no momento do depósito, que inclui o direito de preservação da obra pelo repositório.

O documento final do Plano de Ação para Preservação deve incluir ainda um plano de sucessão com as provisões necessárias para o caso de não haver condições para a continuidade do serviço do repositório Arca.

CONSIDERAÇÕES FINAIS

No decorrer do processo relatado, pudemos confirmar a importância e a relevância das ações de preservação digital, mesmo que discutidas e implementadas em pequena escala e apesar da presença de diversas limitações técnicas e de recursos. Devemos considerar que a preservação digital em um contexto institucional não se trata de uma ação isolada, mas de um processo contínuo de aprendizado e aprimoramento, que vai se construindo gradualmente, e cujos resultados somente serão visíveis e provados em longo prazo.

O resultado dos diagnósticos de confiabilidade e de níveis de preservação são passos intermediários nesse processo, que contribuem para o conhecimento e aperfeiçoamento do repositório no âmbito conceitual e operativo. Concluímos que algumas das recomendações resultantes dos diagnósticos não estão diretamente relacionadas ao Plano de Ação de Preservação Digital, mas podem ser incorporadas ao Plano Operativo do repositório Arca, como parte das atividades do fluxo de trabalho regular, contribuindo para sua melhoria. Entre essas recomendações, destacamos: a adoção de uma rotina de quarentena para verificação antivírus de todos os objetos digitais depositados; o incremento no número de cópias de *backups*, com armazenamento geograficamente distribuído; a criação e manutenção de uma cópia separada dos metadados dos objetos digitais; e o registro documentado das ações de gerenciamento de arquivos para melhorar a capacidade de segurança e a possibilidade de auditar possíveis incidentes. Além disso, é importante estabelecer um plano de recuperação de dados em caso de desastres e comprometimento do sistema.

Ainda que a completa adesão aos padrões internacionais de preservação digital, como o modelo OAIS e os metadados PREMIS, dependam da escolha e implementação de um sistema de preservação digital que contemple recursos mais sofisticados de manutenção de integridade, autenticidade e segurança dos dados, entendemos que a construção de um plano de ação para o

repositório institucional Arca constitui um passo importante tanto para o avanço do debate nacional sobre a preservação digital, como para a consolidação dessas práticas em repositórios institucionais.

Dado o volume de informações relevantes sendo produzido diariamente e a crescente demanda por essas informações, não basta apenas organizar o conhecimento, é preciso preservá-lo a fim de garantir a credibilidade, a confiabilidade e a segurança dos dados em longo prazo, criando políticas e práticas no presente para garantir nossos acervos para o futuro.

REFERÊNCIAS

APREDIG. *Diagnóstico de alcance de recomendaciones NDSA Levels para Fiocruz*. [S.l.:s.n.], 2019.

ARELLANO, M. A. Preservação de documentos digitais. *Ciência da Informação*, Brasília, v. 33, n. 2, p. 15-27, maio/ago. 2004. Disponível em: <http://www.scielo.br/pdf/ci/v33n2/a02v33n2.pdf>. Acesso em: 05 jan. 2019.

ARQUIVO NACIONAL. *Política de Preservação Digital*. [S.l.]: Arquivo Nacional, 2016. Disponível em: http://www.siga.arquivonacional.gov.br/images/an_digital/and_politica_preservacao_digital_v2.pdf. Acesso em: 10 maio 2019.

BECKER, C. *et al.* Systematic planning for digital preservation: evaluating potential strategies and building preservation plans. *International journal on digital libraries*, [S.l.], v. 10, n. 4, p. 133-157, 2009. DOI 10.1007/s00799-009-0057-1. Disponível em: <http://www.ifs.tuwien.ac.at/~becker/pubs/becker-ijdl2009.pdf>. Acesso em: 12 fev. 2019.

BRITISH LIBRARY. *Sustaining The Value: The British Library Digital Preservation Strategy 2017-2020*. [S.l.]: British Library, 2017. Disponível em: https://www.bl.uk/britishlibrary/-/media/bl/global/digital%20preservation/bl_digitalpreservationstrategy_2017-2020.pdf?la=en. Acesso em: 17 jul. 2019.

BULLOCK, A. *Preservation of digital information: issues and current status*. Canada, 22 abr. 1999. Disponível em: <http://www.collectionscanada.ca/9/1/p1-259-e.html>. Acesso em: 28 jul. 2005.

COLOMBIA. MINISTERIO DE EDUCACIÓN NACIONAL. INSTITUTO NACIONAL PARA SORDOS. *Plan de preservación digital*. Colombia: INSOR, 2019. Disponível em: http://www.insor.gov.co/home/wp-content/uploads/filebase/Plan_preservacion_digital_INSOR_2019.pdf. Acesso em: 27 mar. 2019.

- CONSELHO NACIONAL DE ARQUIVOS. *Carta para Preservação do Patrimônio Arquivístico Digital*. Rio de Janeiro: CONARQ, 2005. Disponível em: http://www.conarq.gov.br/images/publicacoes_textos/Carta_preservacao.pdf. Acesso em: 10abr. 2019.
- CORRADO, E. M.; SANDY, H. M. *Digital preservation for libraries, archives and museums*. 2. ed. Lanham, Mariland: Rowman & Littlefield, 2017.
- DIGITAL PRESERVATION COALITION. *Digital Preservation Handbook*. [S.l.], 2015. Disponível em: <https://dpconline.org/handbook>. Acesso em: 07 jan. 2019.
- FIVE COLLEGE CONSORTIUM. *Digital Preservation: A Planning Guide for the Five Colleges*. Amherst, 2017. Disponível em: <https://www.fivecolleges.edu/libraries/digital-preservation/digital-preservation-a-guide-for-the-five-colleges>. Acesso em: 27 fev. 2019.
- FUNDAÇÃO OSWALDO CRUZ. *Política de acesso aberto ao conhecimento*. Rio de Janeiro: Fiocruz/Presidência, 2014. 10 p. Disponível em: https://portal.fiocruz.br/sites/portal.fiocruz.br/files/documentos/portaria_-_politica_de_acesso_aberto_ao_conhecimento_na_fiocruz.pdf. Acesso em: 05 jan. 2019.
- FUNDAÇÃO OSWALDO CRUZ. *Política de preservação dos acervos científicos e culturais da Fiocruz*. Rio de Janeiro: Fiocruz/COC, 2018. 68 p. Disponível em: https://portal.fiocruz.br/sites/portal.fiocruz.br/files/documentos/politica_de_preservacao_dos_acervos_cientificos_e_culturais_da_fiocruz_digital_2018.pdf. Acesso em: 05 jan. 2019.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO/TC 46/SC 11. *Preservación de Documentos Digitales: Guía “Cómo empezar”*. [S.l.:s.n.], 2010. Disponível em: <http://isotc.iso.org/livelink/livelink?func=ll&objId=8800112&objAction=browse&sort=name>. Acesso em: 07 jan. 2019.
- LANZELOTTE, M.F. *et al.* Avaliação de confiabilidade do repositório institucional Arca. In: BIREDIAL-ISTEC, 9., 2019, São Paulo. *Anais* [...]. São Paulo: UNINOVE, 2019. Disponível em: <https://www.arca.fiocruz.br/handle/icict/34552>. Acesso em: 10 set. 2019.
- NATIONAL RECORDS OF SCOTLAND. *The National Records of Scotland and born digital records: a strategy for today and tomorrow*. [S.l.]: National Records of Scotland, 2014. Disponível em: <https://www.nrscotland.gov.uk/files/record-keeping/nrs-digital-preservation-strategy.pdf>. Acesso em: 10 mai. 2019.
- PLANETS. *Preservation Plan Template*. [S.l.:s.n.], 2006. Disponível em: <http://www.ifs.tuwien.ac.at/dp/plato/docs/plan-template.pdf>. Acesso em: 10 abr. 2019.
- PURDUE UNIVERSITY LIBRARIES. PURR *Preservation Strategic Plan*. [S.l.], 2012. Disponível em: <https://purr.purdue.edu/legal/preservation-strategies>. Acesso em: 12 ago. 2019.
- SANTOS, H. M.; FLORES, D. Os impactos da obsolescência tecnológica frente à preservação de documentos digitais. *Brazilian Journal of Information Science: research trends*, [S.l.], v. 11, n. 2, 2017. Disponível em: <http://200.145.171.5/revistas/index.php/bjis/article/view/5550>. Acesso em: 11 mar. 2019.
- THE ROYAL LIBRARY. *Strategy for long term preservation of digital collection materials at the Royal Library*. [S.l.: s.n.], 2014. Disponível em: http://www.kb.dk/export/sites/kb_dk/da/kb/downloadfiler/PreservationStrategyDigitalMaterials-KB-DK-2014.pdf. Acesso em: 10 mai. 2019.
- WEITZEL, S. da R.; MESQUITA, M. A. A. de. Preservação digital em repositórios institucionais: práticas na região Sudeste do Brasil. *Liinc em Revista*, Rio de Janeiro, v. 11, n. 1, p. 181-196, maio 2015. DOI 10.18225/liinc.v11i1.778.